Computer Science and Application 计算机科学与应用, 2025, 15(11), 257-268
Published Online November 2025 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2025.1511302

Hans汉斯

改进YOLOv10n的紧固件检测算法

林俊杰[1]，李　莉[1]，邴志刚[2]

[1]天津职业技术师范大学电子工程学院，天津
[2]天津职业技术师范大学自动化与电气工程学院，天津

收稿日期：2025年10月17日；录用日期：2025年11月17日；发布日期：2025年11月27日

摘　要

针对工业紧固件检测中存在的目标形态多样、堆叠遮挡及背景干扰等问题，提出一种改进的YOLOv10n-WGM检测算法。以M6螺母与M8螺丝为检测目标，基于Unity3D构建多场景仿真数据集，通过控制光照强度、方向、地面纹理与颜色等参数增强数据多样性，并模拟物理掉落过程生成堆叠目标图像，共采集600张图像按4:2划分训练集与验证集。在YOLOv10n基础上引入三方面改进：使用WIoU损失函数提升低质量样本处理能力，通过动态调节损失权重缓解锚框质量不平衡问题，优化梯度分配策略；将C2f模块替换为C2fCIB-LEGM结构，在提升局部特征提取能力的同时建模全局依赖；在检测头前融入CGAFusion模块，提升跨层级特征流动效率并保留空间细节特异性。实验结果表明，YOLOv10n-WGM相比原模型在mAP50、mAP50:95、精确度与召回率分别提升3.588%、1.551%、2.342%与4.295%，漏检率显著降低。该算法具有良好的检测精度与鲁棒性，适用于工业环境中的紧固件实时检测需求。

关键词

紧固件检测，YOLOv10n，WIoU，LEGM，CGA Fusion

An Improved YOLOv10n Algorithm for Industrial Fastener Detection

Junjie Lin[1], Li Li[1], Zhigang Bing[2]

[1]School of Electronic Engineering, Tianjin University of Technology and Education, Tianjin
[2]School of Automation Electrical Engineering, Tianjin University of Technology and Education, Tianjin

Received: October 17, 2025; accepted: November 17, 2025; published: November 27, 2025

Abstract

To tackle challenges like varied shapes, occlusion, and background clutter in industrial fastener detection, we propose an enhanced YOLOv10n-WGM algorithm. Focusing on M6 nuts and M8 screws, a diverse synthetic dataset was generated using Unity3D by varying lighting, textures, and colors,

文章引用: 林俊杰, 李莉, 邴志刚. 改进 YOLOv10n 的紧固件检测算法[J]. 计算机科学与应用, 2025, 15(11): 257-268.
DOI: 10.12677/csa.2025.1511302

and simulating physical stacking. A total of 600 images were collected and split into training and validation sets at a 4:2 ratio. The model incorporates three key improvements: 1) WIoU loss, which dynamically adjusts the loss weight to alleviate anchor box quality imbalance and optimize the gradient allocation strategy, for better handling of low-quality samples; 2) The C2f module is replaced by the C2fCIB-LEGM structure to enhance local feature extraction and model global dependencies; and 3) The CGAFusion module is integrated before the detection head to enhance cross-level feature flow efficiency and preserve spatial detail specificity. Experiments show significant gains: +3.588% mAP50, +1.551% mAP50:95, +2.342% precision, and +4.295% recall compared to the original model, with markedly lower missed detections. The method demonstrates high accuracy and robustness for real-time industrial fastener inspection.

Keywords

Industrial Fastener Detection, YOLOv10n, WIoU, LEGM, CGA Fusion

1. 引言

针对以 M8 螺丝、M6 螺母为代表的工业常用紧固件的视觉抓取任务在智能制造领域具有重要意义，广泛应用于汽车装配[1]、航空航天[2]与重型机械制造[3]等工业场景。在这些应用场景中，视觉环境光照条件常常变化，特别是在大量堆叠的紧固件检测任务中，目标彼此遮挡、类内差异小以及实时性要求高等因素常，带来多重挑战，导致抓取任务失败。传统的视觉检测方法多依赖于边缘特征提取和模板匹配，但在堆叠遮挡和复杂光照变化下，特征区分能力有限，无法保证检测精度。文献[4]提出基于注意力机制的多尺度检测网络，通过引入通道－空间双重注意力模块，显著提升遮挡目标的识别能力。但该网络计算时间复杂度较高，单帧推理耗时 3.8 秒，无法满足工业生产线视频帧实时检测需求；文献[5]则创新性地融合 Transformer 与卷积架构，构建了 FastenerFormer 模型，在复杂工业场景中达到 92.5%的检测精度。但该模型在非常规光照条件下的紧固件检测任务中性能不稳定，在过曝光场景下的误检率超过 15%；文献[6]提出改进 YOLOv8 的工业紧固件检测算法，通过引入可变形卷积和自适应特征融合策略，有效提升堆叠场景中遮挡目标的检测精度。但在高度重叠情况下仍存在约 18%的漏检率。

特征金字塔网络在堆叠目标检测任务中通常起到多尺度特征融合的作用；文献[7]提出动态特征金字塔 DFPN，通过权重学习机制自适应融合多尺度特征，在复杂工业检测场景有一定性能提升。但是 DFPN 没有处理不同层次特征间的语义差异，直接融合会导致特征冲突，在一些紧固件数据集上的检测精度低；文献[8]提出基于注意力引导的特征融合 AGFF，其能够根据特征重要性动态调整融合权重，相比常规 FPN 提升特征融合效果。然而 AGFF 模块因其复杂的注意力计算，导致推理速度下降 20%，因此难以满足工业实时检测需求。

随着对比学习技术的发展，文献[9]提出轻量级特征提取网络 LightNet，通过深度可分离卷积和通道 shuffle 操作实现高效特征提取。LightNet 参数量仅为 ResNet-50 的 1/8，推理速度提升 5 倍，成为工业检测部署的参考架构。然而，LightNet 模型在高度遮挡场景下的鲁棒性不足，对堆叠目标的区分度不够，造成检测性能下降。

针对堆叠目标检测难题，文献[10]设计解耦检测头网络 DDH-Net，通过分类与回归任务分离和特征细化机制，将紧固件堆叠场景中的召回率提升至 75.6%，但模型对光照变化的适应性仍有待提高。

尽管已有大量研究致力于提升工业场景背景下的目标检测任务的精度，但现有方法在堆叠目标检测特殊环境下表现依旧存在不足。针对以上问题，本文提出了一种基于 YOLOv10n 的改进算法 YOLOv10n-WGM。该算法引入了加权梯度调制机制(WGM)和内容感知的空间通道注意力模块，在不影响模型轻量化的基础上，提升了堆叠紧固件环境的检测精度和召回率，显著增强了模型算法的鲁棒性。该算法为工业堆叠紧固件视觉检测提供了一种高效、可靠的解决方案。

2. YOLOv10n 算法改进

2.1. YOLOv10 简述与改进算法

YOLOv10 是由清华大学研究团队于 2024 年 5 月发布的端到端目标检测算法[11]，在检测精度、推理速度与计算效率方面较上一代均有显著提升。该模型引入跨阶段局部网络(CSPNet)并结合自适应通道分配机制，依据目标图像特征优化分支通道数以减少计算冗余；同时借助跨层特征跳跃连接与模型剪枝策略，将参数量压缩至 YOLOv9 的 78%。YOLOv10 根据目标设备性能考量划分为多个版本，其中 YOLOv10n 为轻量化版本，通过网络剪枝、降低参数量和计算复杂度，从而减小模型体积并提升推理速度。在 NVIDIA T4 TensorRT 平台测试中，YOLOv10n 实现了 1.56 ms/帧的推理速度，较 YOLOv8n 与 YOLOv9n 提升近 7 倍，尤其适用于资源受限环境下的实时检测任务。

尽管 YOLOv10 已具备优良性能，但在堆叠环境下的紧固件目标检测任务中，其检测精度与速度仍面临挑战。为此，本文以 YOLOv10n 为基准模型，提出一种适用于紧固件堆叠场景的改进模型 YOLOv10n-WGM (结构如图 1 所示)，旨在进一步提升模型的检测能力。

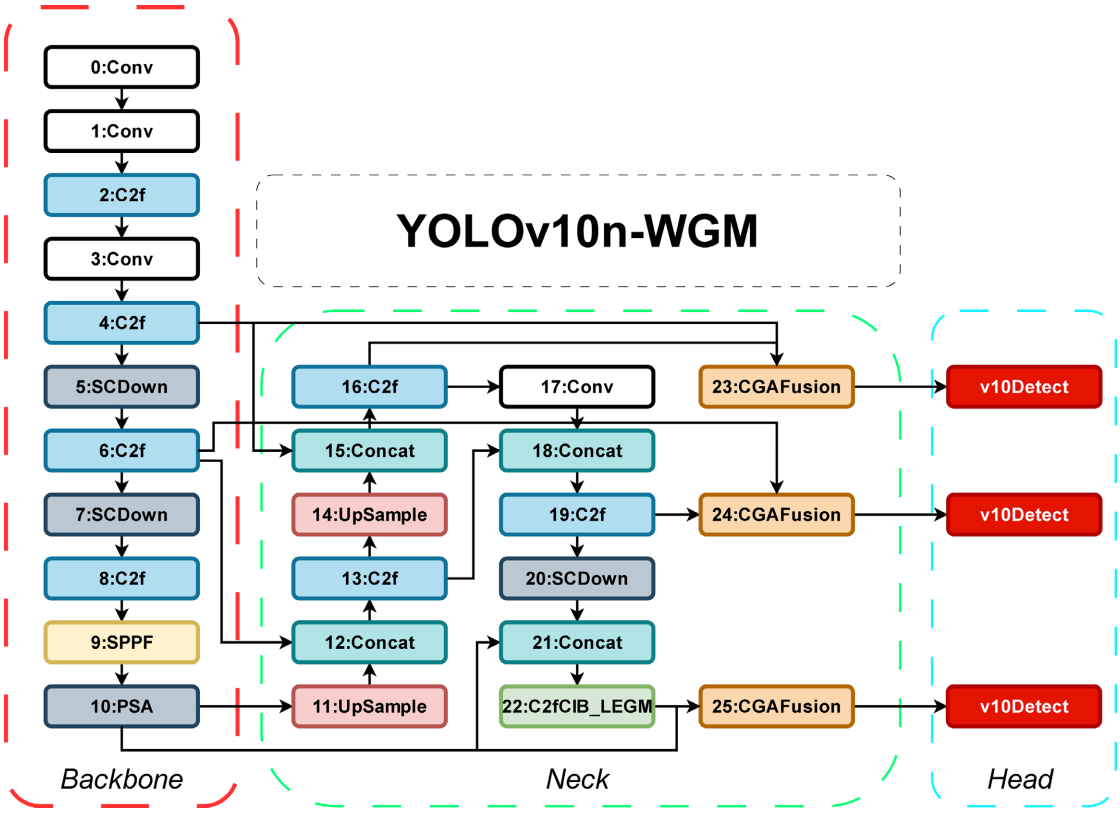


Figure 1. YOLOv10n-WGM model architecture
图 1. YOLOv10n-WGM 模型结构图

2.2. LEGM 局部特征嵌入全局特征提取模块

本文实验所用数据集由于包含大量强光、阴影以及明暗交替等多种极端光照条件下的样本，即所谓高动态范围(HDR)图像，面临局部过曝或欠曝等干扰因素导致的细节丢失、伪影以及目标特征难以捕捉的问题。传统卷积网络(如 ResNet、DarkNet)对于 HDR 图像的极端光照处理能力不足。而专门的 HDR 重建网络(如 HDRNet、HDRfeat)存在预处理流程，不适用于实时目标检测。此外基于注意力机制的特征提取模块(如 SENet、CBAM)未能考虑 HDR 图像的动态范围问题带来额外的计算复杂度。本文针对以上问题引入了一种称为局部增强与全局调制网络(LEGM) [12]的结构。该网络的整体架构如图 2 所示。

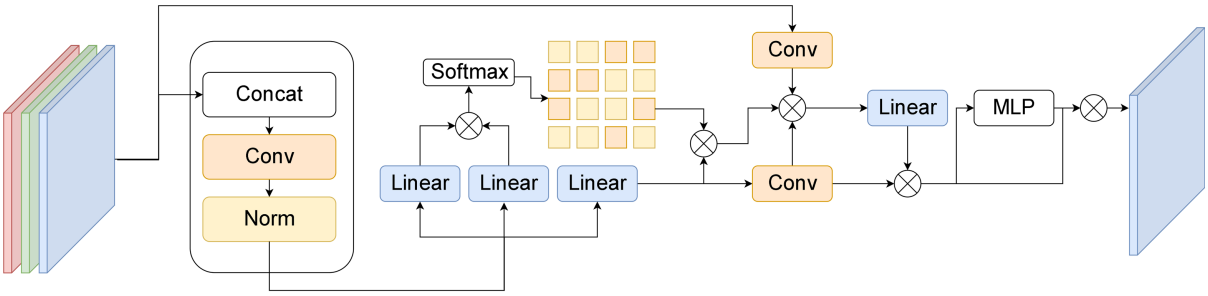


Figure 2. Schematic diagram of LEGM principle
图 2. LEGM 原理示意图

LEGM 的核心由两个分支组成：1) 局部增强分支(Local Enhancement Branch, LEB)：该分支专注于弥补 HDR 图像中因极端光照干扰而缺失的局部细节。它采用可学习的局部特征提取模块，结合可变形卷积(Deformable Convolution)和局部注意力机制，自适应聚焦于潜在目标区域，增强微小目标和遮挡目标的边缘及纹理特征表示；2) 全局调制分支(Global Modulation Branch, GMB)：该分支捕捉图像的全局上下文信息，并调节特征表示以适应 HDR 场景中的广泛动态范围。它利用全局平均池化和频域变换(如傅里叶变换)提取全局特征，并融合了门控机制与局部特征，从而增强模型对复杂光照条件的鲁棒性。

LEGM 特征图输入大小为 $C \times H \times W$ (其中 C 为通道维度，H 和 W 分别为空间高度和宽度)，首先在局部增强分支中进行细节恢复，输出增强后的特征；同时在全局调制分支中提取全局上下文信息。两个分支的特征通过自适应融合模块进行整合，最终输出优化后的特征图。

不同于以往模型依赖多曝光序列或复杂 HDR 重建，LEGM 直接处理单张 HDR 图像(或 SDR 转 HDR 后的图像)，在单一网络中通过双分支结构实现局部细节恢复和全局上下文调制，不需要额外的预处理步骤或复杂的合成操作，显著降低计算开销。此外，LEGM 仅使用普通卷积和常见注意力机制，不依赖额外的 CUDA 算子，在各种深度学习框架(如 PyTorch)中都能实现和部署，尤其适用于计算资源受限的移动端或嵌入式设备等边缘设备上的实时 HDR 图像目标检测任务。

2.3. CGA Fusion 内容引导的注意力融合模块

在复杂工业场景的目标检测任务中，模型需要准确识别被遮挡或隐藏等因素干扰的多尺度目标。传统方法在处理此类场景时表现不佳：固定尺寸的卷积核难以自适应地聚焦于不同尺度目标的关键特征区域，直接扩大感受野又会带来计算复杂度急剧上升的问题。在目标密集堆叠的情况下，相邻目标的特征提取可能导致特征混淆和识别精度下降。零部件(如螺栓、螺母)往往以高密度方式堆积，存在严重的相互遮挡和部分可见情况。传统的卷积操作采用统一的处理方式，无法根据目标的空间分布和语义重要性进行差异化特征提取。特征金字塔网络(FPN, Feature Pyramid Network)采用垂直路径和横向连接的方式融合多尺度特征，但在密集遮挡下效果有限。路径聚合网络(PANet, Path Aggregation Network)在原有 FPN 的

基础上将单向的垂直路径改为双向，进一步增强了特征融合，但仍未能实现内容自适应，缺少对特征语义重要性的考虑。

CGA Fusion (Content-Guided Attention Fusion) [13]创新性地提出了一种基于内容感知的特征优化机制，其核心在于通过一个轻量的子网络，动态生成与输入特征内容相关的空间与通道注意力权重，从而实现对关键特征的自适应增强与对背景干扰的有效抑制。该模块的整体架构如图 3 所示。

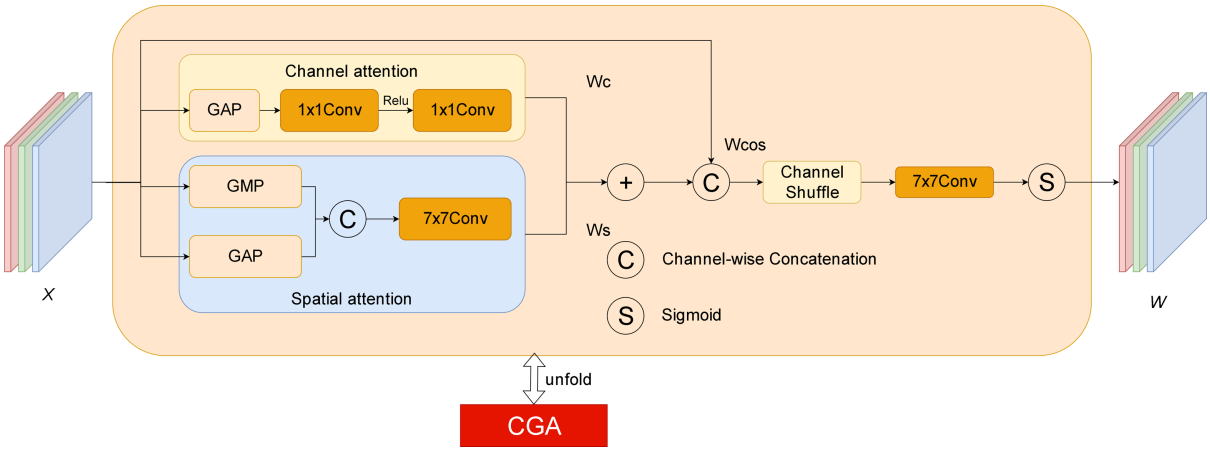


Figure 3. Schematic diagram of CGA Fusion principle
图 3. CGA Fusion 原理示意图

其核心运算过程可由以下公式组定义：

$$\boldsymbol{X} \in \mathbb{R}^{C\times H\times W} \tag{1}$$

$$\boldsymbol{K} = \mathrm{Conv1x1\_k}\left(\boldsymbol{X}\right) \tag{2}$$

$$\boldsymbol{Q} = \mathrm{Conv1x1\_q}\left(\boldsymbol{X}\right) \tag{3}$$

$$\boldsymbol{A}_{\mathrm{raw}} = \mathrm{Softmax}\left(\mathrm{Reshape}\left(\boldsymbol{Q}\right)^{\top}\ \mathrm{Reshape}\left(\boldsymbol{K}\right)\right) \tag{4}$$

$$\gamma = \sigma\left(\mathrm{FC}_2\left(\delta\left(\mathrm{FC}_1\left(\mathrm{GAP}\left(\boldsymbol{X}\right)\right)\right)\right)\right) \tag{5}$$

公式 1 中：$\boldsymbol{X}$ 为给定输入特征图，公式(2)和公式(3)通过两个独立的 1x1 卷积层生成键(Key)和查询(Query)特征图。随后，公式(4)通过计算查询和键的相似性矩阵来获取全局上下文信息，生成初始的空间注意力蓝图 $\boldsymbol{A}_{\mathrm{raw}}$，其中 Softmax 沿矩阵行方向进行归一化，Reshape 操作将特征图的空间维度展平。同时，公式(5)利用全局平均池化(GAP)捕获通道级统计信息，并通过包含全连接层和激活函数的门控机制生成通道注意力权重。最终的空间注意力权重图由初始蓝图 $\boldsymbol{A}_{\mathrm{raw}}$ 经过一个卷积层细化后，与通道权重 γ 进行元素相乘，实现空间与通道的协同调制。

2.4. Wise-IoU 损失函数

YOLOv10n 中的回归损失函数 CIoU 考虑了边界框的中心距离、宽高比及交并比，在单个目标以及尺度较为统一的目标检测任务中性能稳定，但面对目标堆叠、遮挡、及类别不平衡等问题时漏检率高，效果大打折扣。CIoU 中的宽高比惩罚项未能考虑数据集的样本质量问题，因此一些低质量样本可能影响优化方向。EIoU 直接最小化宽高差，收敛速度和定位精度超过了 CIoU，但其仍未完全解决低质量样本的梯度分配问题。Alpha-IoU 尝试通过调整超参数来改变损失函数的形状，但对于不同数据集都需要进行反复调整，未能从根本上解决此类问题。

为解决以上问题，本文提出采用 Wise-IoU (WIoU) [14]损失函数，WIoU 从评估样本质量的角度看待样本宽高比不一致和梯度消失的问题。在 WIoU 中离群度(Outlierness)代表样本的质量等级，例如将扁平锚框等有着奇怪宽高比的样本视为高离群度样本，适当增加该类样本的梯度。而对于低离群度，也就是宽高比较为正常的样本则适当减少梯度。此设计不仅避免设置专门的宽高比惩罚项，而且保证梯度始终稳定有效。通过引入离群度聚焦机制，显著提升模型在堆叠场景下的检测性能。其通过评估边界框的质量来自适应调整梯度增益，使模型更加关注难以学习的样本(如严重遮挡的螺丝螺母)，同时减少对简单样本的过度关注。Wise-IoU 的计算公式如下：

$$\mathcal{L}_{WIoU} = R_{WIoU} \cdot L_{IoU} \tag{6}$$

$$R_{WIoU} = \exp\left(\frac{\left(x - x_{gt}\right)^2 + \left(y - y_{gt}\right)^2}{\left(W_g^2 + H_g^2\right)^*}\right) \tag{7}$$

$$L_{IoU} = 1 - \frac{\left|B \cap B_{gt}\right|}{\left|B \cup B_{gt}\right|} \tag{8}$$

公式(7)中 x、y 为预测框中心坐标，x_{gt}、y_{gt} 为真实框坐标。W_g^2、H_g^2 为最小外接矩形的宽和高，该公式通过区分预测框质量来为不同预测框分配不同权重。公式(8)中 B 和 B_{gt} 分别表示预测框和真实框的区域，L_{IoU} 计算基础的 IoU 损失项，L_{IoU} 越趋近于 0，代表预测框和真实框的重合程度越高，定位越精准。最后在公式(6)中为不同预测框的易学习程度进行动态加权并得到最终损失。

与传统的 CIoU 损失相比，WIoU 损失在保持大尺寸零件检测性能的同时，显著提升了小尺寸螺母和部分遮挡螺丝的定位准确性。其动态调节机制特别适用于工业场景中常见的零件堆叠、相互遮挡等复杂情况，在提高检测精度的同时保持了模型的收敛稳定性，为紧固件的自动化检测提供了有效的解决方案。

3. 实验设计与结果分析

3.1. 数据集构建

实验基于 Unity3D 仿真引擎搭建了数据采集平台，以该平台为基础构建了面向紧固件检测的自制数据集。该平台模拟多种表面纹理(如布料、石料)，并实时调节方向光源的位置与旋转，生成多张目标处于光照或阴影环境下的图像；光强参数随时间变化波动，以模拟强光与弱光等不同照明条件下的视觉干扰。实验以 M6 螺母和 M8 螺丝两类常见紧固件作为检测目标，每种型号分别提供黑色与银色两种外观。利用该平台，采集了 M6 和 M8 紧固件的单独图像(包含黑、银两种外观)各 200 张，此外还生成了 200 张包含多个螺母与螺丝混合堆叠场景的图像。图像生成过程中，目标在平台上方随机位置生成后，开启物理引擎模拟自由重力与碰撞，确保每次采集的图像在布局和遮挡状态上都具有差异性。最终共获得 600 张图像，将 400 张单独目标图像作为训练集，200 张混合堆叠图像作为验证集。此划分策略旨在通过验证集专门评估不同模型在复杂堆叠场景中的目标识别与抗遮挡性能。数据集部分图像如图 4 所示。

3.2. 实验环境与评估指标

3.2.1. 实验环境

操作系统为 Ubuntu 20.04 LTS，配备 16 GB 内存和 RTX 3060 (12 GB 显存)。深度学习框架为 Pytorch 2.6.0，通过 CUDA 12.6 进行 GPU 加速。训练参数包含训练轮数 400、图像长放缩到 640，宽自动补全，输入尺寸为 640 × 640、Batch Size 15、初始学习率 0.01、动量大小 0.937 和权重衰减系数 0.0005。

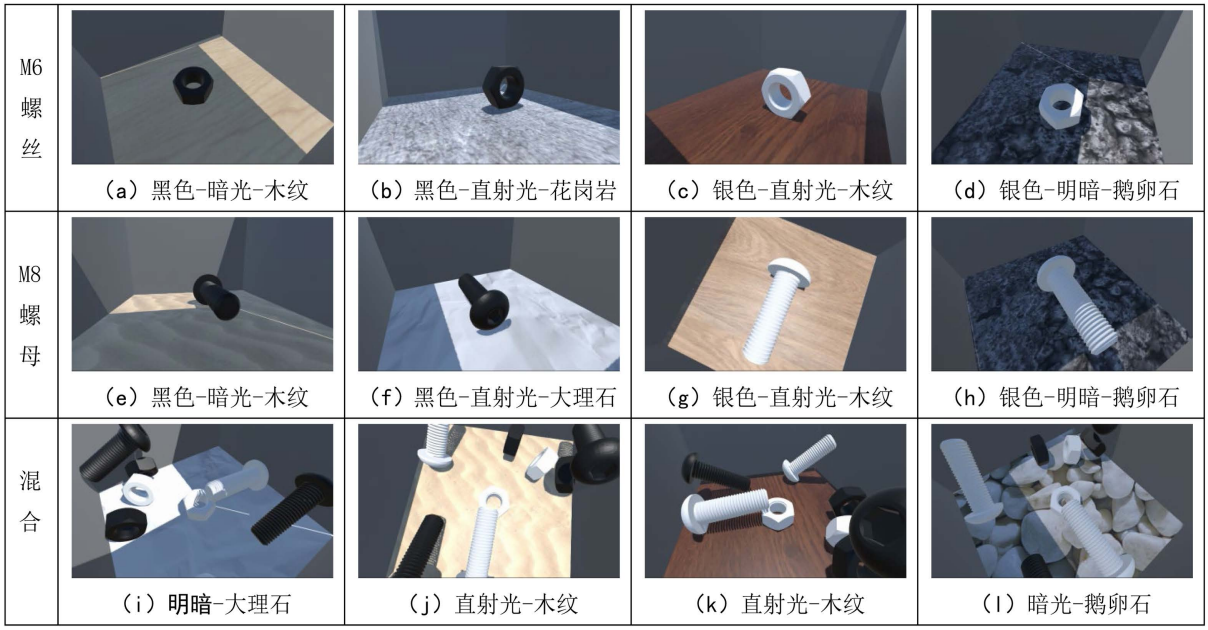


Figure 4. Schematic diagram of the self-created dataset
图 4. 自制数据集示意图

3.2.2. 评估指标

为了评估模型的整体性能，本实验使用的评估指标为平均精度均值(mAP)、精度(Precision P)召回率(Recall R)、浮点运算数(Floating-point Operations Per Second)、模型大小和参数量。mAP 计算公式为

$$\mathrm{mAP}=\frac{1}{N}\sum_{c=1}^{N}\mathrm{AP}_c \tag{9}$$

公式(9)中，mAP 反映了模型在所有类别中的检测精度和召回率之间的平衡(单位：%)，mAP@0.5 是指当 IoU 为 0.5 时的 mAP 值； AP_c 为第 c 类别的平均精度；n 为类别总数。

3.3. 实验结果分析

3.3.1. 消融实验分析

为了进一步分析验证每一个改进对航拍图像检测性能的有效性，本文以 YOLOv10n 为基准模型进行一系列消融实验，实验结果如表 1 所示。

Table 1. Ablation experiment results
表 1. 消融实验结果

基准模型	WIoU	LEGM	CGA Fusion	mAP@0.5/%	P/%	R/%	大小/MB	Layer /层数	参数量 /MB	浮点运算 /GFlops
YOLOv10n				61.75	79.34	49.36	5.50	102	2.265	6.5
	√			61.38	77.9	49.4	5.5	102	2.265	6.5
	√	√		60.95	76.11	51.6	5.57	125	2.292	6.6
	√		√	67.838	80.52	55.25	5.81	129	2.418	6.8
	√	√	√	65.34	81.68	53.66	6.04	139	2.539	6.9

根据表 1 结果可知：首先，单独引入 WIoU 损失函数时，模型参数量和计算量保持不变，但 mAP@0.5 有略微下降，这表明直接替换损失函数未能带来性能增益，结合 WIoU 函数特点分析，数据集此时被判定为高质量样本即简单样本占比可能更多，由于损失函数特性反而削弱了重要学习信号，因此得出其需要与其他模块协同工作的结论；其次，当同时采用 WIoU 和 LEGM 模块时，mAP@0.5 下降至 60.95%，但召回率提升了 2.24%，同时模型复杂度和计算开销仅有小幅增加，证明 LEGM 模块在不影响参数轻量化的同时增强了对困难样本的检测能力，LEGM 作为特征提取模块为样本带来了更丰富、复杂的特征表示，弥补了原来的简单样本缺陷，这使得 WIoU 重新评估样本的质量等级；再次，当结合 WIoU 与 CGA Fusion 模块时，mAP@0.5 显著提升 6.088%，精确度和召回率分别提高 1.18%和 5.89%，以少量增加模型大小和计算量为代价，带来显著性能大幅度提升，表明 CGA Fusion 模块通过增强特征融合能力能极大克服堆叠目标样本带来的特征混淆问题，也进一步优化 WIoU 对于样本的权重分配策略，改善了模型性能；最后，当同时集成所有三个改进模块时，mAP@0.5 达到 65.34%，较基准提升 3.59%，精确度和召回率分别提高 2.34%和 4.3%，而模型复杂度和计算开销控制在合理范围内。

上述消融实验表明，本文提出的各改进模块对紧固件堆叠目标检测各有独特贡献：WIoU 损失函数为模型优化提供更好的回归准则；LEGM 模块以较低的计算代价提升了对困难样本的敏感性；CGA Fusion 模块通过高效的特征融合显著提升了检测精度。LEGM 模块与 CGA Fusion 模块通过各自的特征增强策略加强了对原有样本的特征表示能力，避免 WIoU 损失函数过早丢失部分优质样本，对数据集的兼容性更好。三个模块在保持模型轻量化的同时，有效提升了模型的检测性能。

3.3.2. 对比实验分析

为验证本文提出改进算法的有效性，将本文算法与其他该领域的常见算法进行比对分析，结果如表 2 所示。

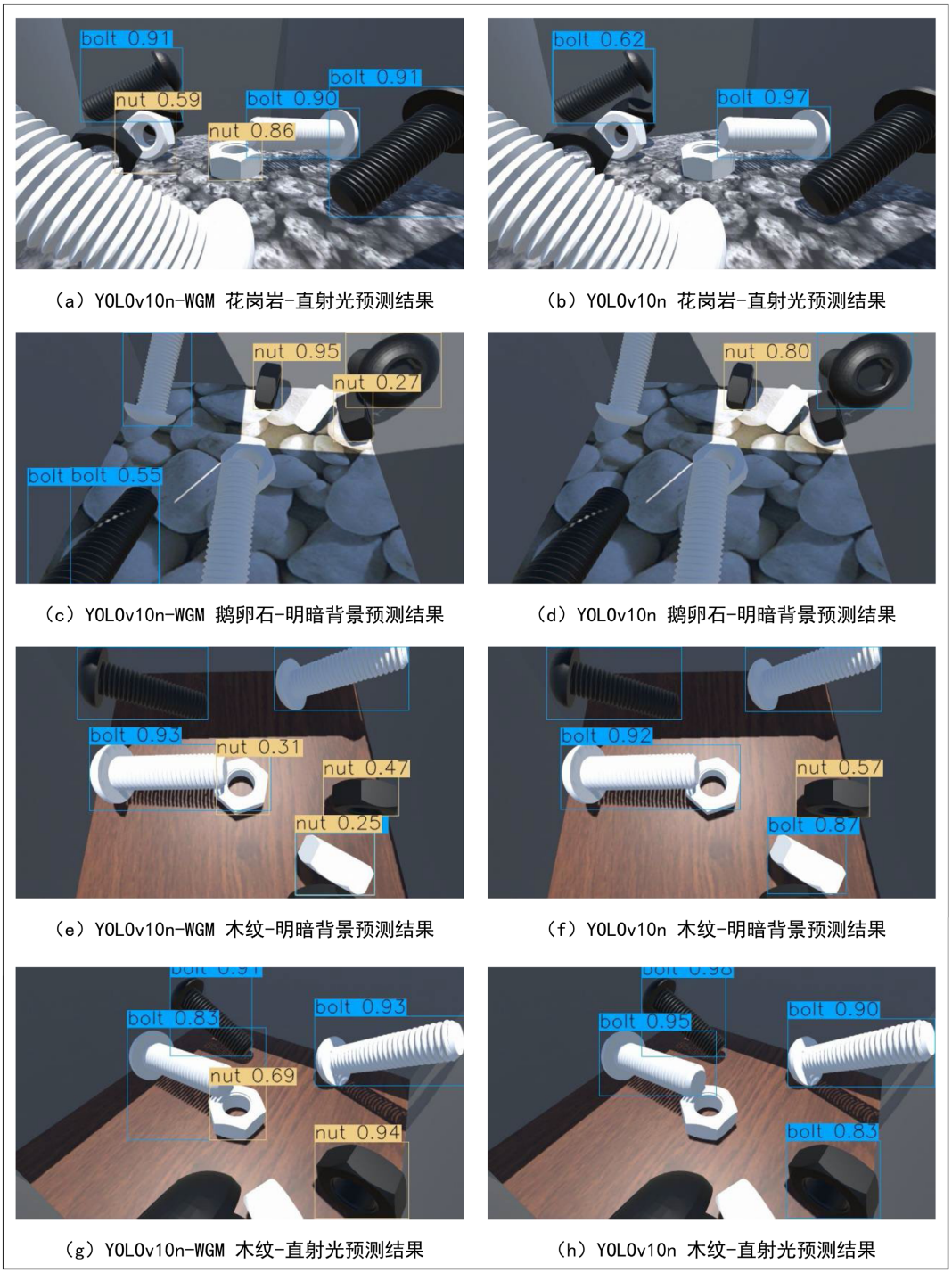
Table 2. Comparative experimental results of various detection models
表 2. 各种检测模型对比实验结果

方法	mAP@0.5/%	*P*%	*R*%	大小/MB	参数量/MB	浮点运算/GFlops
YOLOv5n	70.99	82.28	59.38	5.08	2.50	7.1
YOLOv6n	70.56	77.28	60.96	8.30	4.233	11.7
YOLOv8n	71.19	82.65	58.89	5.96	3.006	8.1
YOLOv8n_ghost	57.91	74.11	53.17	3.59	1.714	5.0
YOLOv9t	60.54	77.75	51.32	4.44	1.971	7.6
YOLOv10n	61.75	79.34	49.36	5.50	2.266	6.5
YOLOv10n-WGM	65.34	81.68	53.66	6.04	2.539	6.9

表 2 结果可知：YOLOv10n-WGM 在 mAP@0.5 即综合性能上比 YOLOv10n、YOLOv9t、YOLOv8-ghost 等模型分别高出 3.59%、4.8%、7.43%，排在对比模型中前列；YOLOv10n-WGM 的精度 P 则比 YOLOv6n、YOLOv8n-ghost、YOLOv9t、YOLOv10n 等模型分别高出 4.4%、7.57%、3.93%、2.34%。YOLOv10n-WGM 在综合性能虽稍逊于 YOLOv5n、YOLOv6n、YOLOv8n 等模型，但其浮点运算量比这三种模型更低，模型参数量则比 YOLOv6n、YOLOv8n 更低，在边缘计算设备上更具优势。

3.3.3. 可视化对比分析

为验证本文算法(YOLOv10n-WGM)在实际背景中的性能，与基准模型 YOLOv10n 进行了可视化对比。实验选取了四组代表图像，涵盖明暗背景和直射光场景，每组包含 YOLOv10n 和 YOLOv10n-WGM 的预测结果，对比结果如图 5 所示。



（a）YOLOv10n-WGM 花岗岩-直射光预测结果　（b）YOLOv10n 花岗岩-直射光预测结果

（c）YOLOv10n-WGM 鹅卵石-明暗背景预测结果　（d）YOLOv10n 鹅卵石-明暗背景预测结果

（e）YOLOv10n-WGM 木纹-明暗背景预测结果　（f）YOLOv10n 木纹-明暗背景预测结果

（g）YOLOv10n-WGM 木纹-直射光预测结果　（h）YOLOv10n 木纹-直射光预测结果

Figure 5. Detection performance comparison between YOLOv10n-WGM and YOLOv10n
图 5. YOLOv10n-WGM 和 YOLOv10n 在 4 种不同场景中检测效果对比图

在一、三、四组图像对比中，无论是强烈明暗对比还是直射光下，YOLOv10n-WGM 均能识别出图像中心被螺丝部分遮挡的白色 M6 螺母，有效减少漏检，尤其在目标与背景相似或纹理复杂时表现更佳，证明其在处理光照变化和背景干扰方面具有优势；在第二组图像对比中，面对右上角强烈直射光导致的过曝或左侧阴影过深部分，YOLOv10n-WGM 能够识别到左侧的 M8 螺丝目标以及右上角的 M6 螺母识别置信度更高，依然保持高识别准确率；在三、四组图像对比中，右下角的 M6 螺母目标均被 YOLOv10n 误识别为 Bolts，而反观 YOLOv10n-WGM 则能正确识别，有效减少误检，从而提升模型精度，凸显其在应对复杂光照环境方面的优越性；综合对比表明，本文算法 YOLOv10n-WGM 在明暗背景和直射光等复杂环境下，均显著优于基准模型 YOLOv10n，提供了更准确、全面的目标识别结果，验证了其在实际应用中的有效性和鲁棒性。

此外，为了更直观地看到本文算法和基准模型的评估指标变化情况，图 6~8 依次展示了算法精度、召回率、mAP@0.5 等模型收敛对比图，可以看到模型在 400 个 epoch 时趋近收敛。图 9 展示了模型收敛后的精度、召回率、mAP@0.5 等数据的柱状对比图，可以看到本文算法模型精度、召回率、mAP@0.5 均高于基准模型，验证本文的 YOLOv10n-WGM 具有明显的优势。

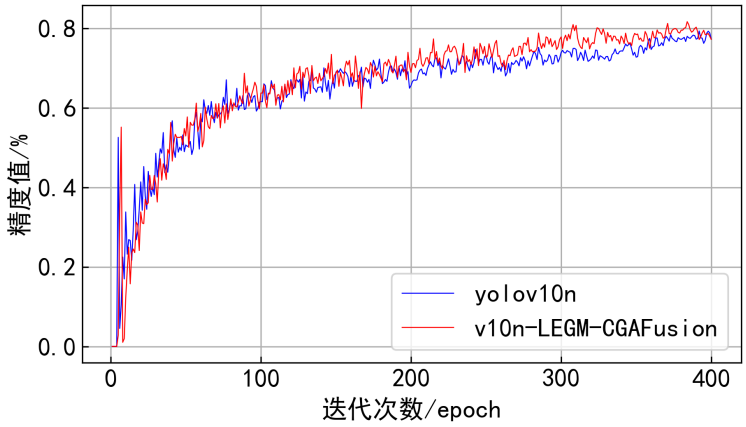


Figure 6. Comparison of accuracy metrics
图 6. 精度指标对比图

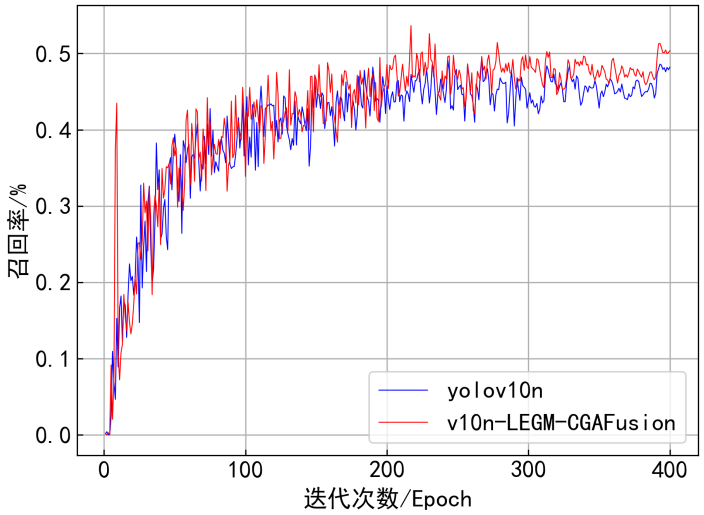


Figure 7. Recall metric comparison
图 7. 召回率指标对比图

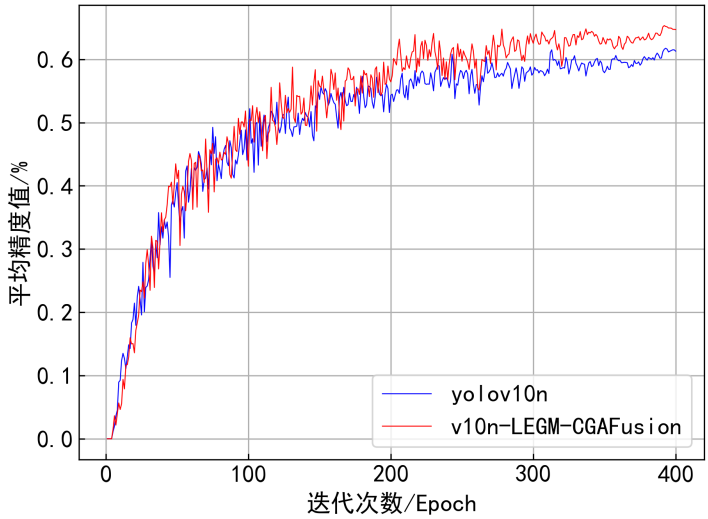


Figure 8. Comparison of mAP@0.5
图 8. mAP@0.5 指标对比图

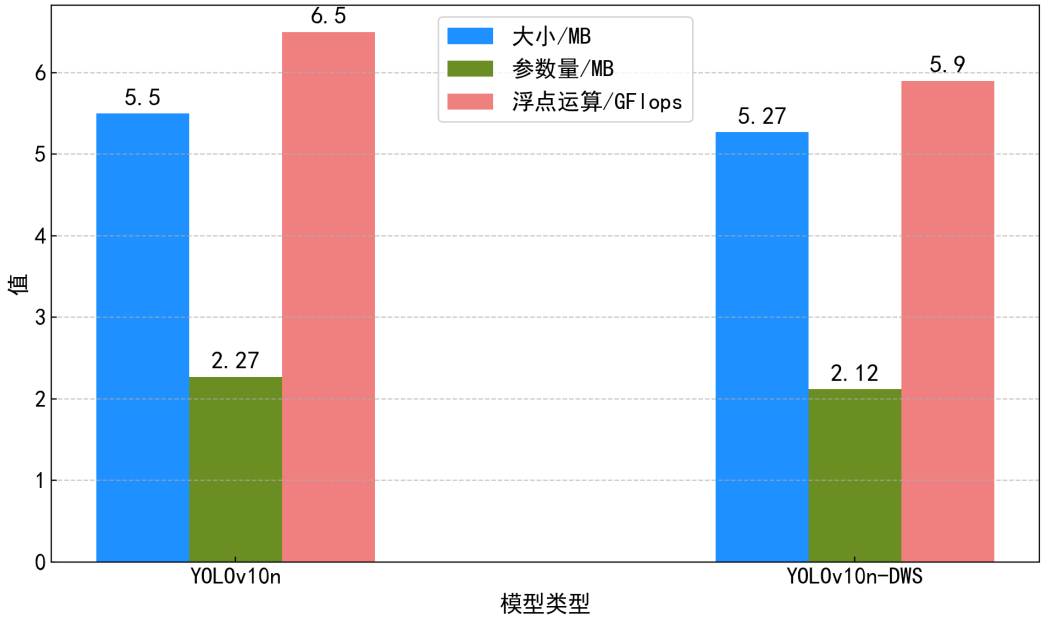


Figure 9. Bar chart comparison of accuracy, recall and mAP@0.5
图 9. 精度、召回率、mAP@0.5 柱状对比图

4. 结语

本文针对工业紧固件检测中的挑战，提出了一种改进的 YOLOv10n-WGM 检测算法。该算法在 YOLOv10n 基础上，通过引入 Wise-IoU 损失函数优化低质量样本处理，C2fCIB-LEGM 结构增强局部特征与全局依赖建模，以及 CGAFusion 模块提升跨层级特征融合效率和抑制背景干扰。实验结果表明，YOLOv10n-WGM 在自制数据集上，mAP50、mAP50:95、精确度与召回率均有显著提升，漏检率降低。该算法在复杂工业环境下展现出良好的检测精度与鲁棒性，为紧固件实时检测提供了高效可靠的解决方案。未来工作将继续优化模型推理速度，探索更轻量级的注意力机制和特征融合策略，并研究半监督或无监督学习方法，以提升算法在实际工业场景中的泛化能力和应用潜力。

参考文献

[1] 代金龙, 习吕鹏, 王亮, 等. 商用车螺纹紧固件连接防松技术分析及理论研究[J/OL]. 汽车工艺与材料, 1-7. https://doi.org/10.19710/J.cnki.1003-8817.20250162, 2025-09-15.

[2] 常星星, 陈卫林, 黄超, 等. 低合金超高强度钢(30CrMnSiNi2A)螺栓结构件的加工阐述[J]. 内江科技, 2025, 46(8): 75-76.

[3] 张文健. 新型钛合金机械紧固件锻造温度优化方案[J]. 锻压装备与制造技术, 2025, 60(4): 104-107.

[4] 陈卫彪, 贾小军, 朱响斌, 冉二飞, 魏远旺. G-YOLOv7: 面向无人机航拍图像的目标检测算法[J]. 光电子·激光, 2025, 36(2): 146-157.

[5] Zhu, X., Lyu, S., Wang, X. and Zhao, Q. (2021) TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. 2021 *IEEE/CVF International Conference on Computer Vision Workshops* (*ICCVW*), Montreal, 11-17 October 2021, 2778-2788. https://doi.org/10.1109/iccvw54120.2021.00312

[6] 王晓辉, 贾韫硕, 郭丰娟. 基于改进 YOLOv8 的汽车门板紧固件检测算法[J]. 计算机工程与设计, 2025, 46(1): 298-306.

[7] Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A.P., Bishop, R., *et al.* (2016) Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 1874-1883. https://doi.org/10.1109/cvpr.2016.207

[8] Wang, J., Chen, K., Xu, R., Liu, Z., Loy, C.C. and Lin, D. (2019) CARAFE: Content-Aware Reassembly of Features. 2019 *IEEE/CVF International Conference on Computer Vision* (*ICCV*), Seoul, 27 October-2 November 2019, 3007-3016. https://doi.org/10.1109/iccv.2019.00310

[9] Ye, C., Zhao, C., Yang, Y., Fermüller, C. and Aloimonos, Y. (2016) LightNet: A Versatile, Standalone Matlab-Based Environment for Deep Learning. *Proceedings of the* 24*th ACM international conference on Multimedia*, Amsterdam, 15-19 October 2016, 1156-1159. https://doi.org/10.1145/2964284.2973791

[10] Yu, J., Mao, J., Wang, Y., He, Z., Yi, J., Tao, Z., *et al.* (2025) DDH-Net: A Dual Dynamic Head Network for Transmission Line Inspection. *IEEE Transactions on Instrumentation and Measurement*. https://doi.org/10.1109/tim.2025.3529547

[11] Wang, A., *et al.* (2024) YOLOv10: Real-Time End-to-End Object Detection. arXiv: 2405.14458

[12] Zhang, Y., Zhou, S. and Li, H. (2024) Depth Information Assisted Collaborative Mutual Promotion Network for Single Image Dehazing. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Seattle, 16-22 June 2024, 2846-2855. https://doi.org/10.1109/cvpr52733.2024.00275

[13] Chen, Z., He, Z. and Lu, Z. (2024) DEA-Net: Single Image Dehazing Based on Detail-Enhanced Convolution and Content-Guided Attention. *IEEE Transactions on Image Processing*, **33**, 1002-1015. https://doi.org/10.1109/tip.2024.3354108

[14] Tong, Z., Chen, Y., Xu, Z., *et al.* (2023) Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. arXiv: 2301.10051